



Robust pedestrian detection in thermal infrared imagery using the wavelet transform

Jianfu Li, Weiguo Gong*, Weihong Li, Xiaoying Liu

Key Laboratory for Optoelectronic Technology and Systems of the Education Ministry of China, Chongqing University, Chongqing 400044, China

ARTICLE INFO

Article history:

Received 20 November 2009

Available online 8 April 2010

Keywords:

Pedestrian detection

Thermal infrared imagery

Double-density dual-tree complex wavelet transform (DD-DT CWT)

Support vector machine (SVM)

ABSTRACT

A novel and robust pedestrian detection method in thermal infrared images based on the double-density dual-tree complex wavelet transform (DD-DT CWT) and wavelet entropy is presented in this paper. The regions of interest (ROIs) are located first making use of high brightness property of the pedestrian pixels caused by the self-emission of the pedestrians related to the Planck's law. The candidate ROIs are then decomposed by DD-DT CWT and the wavelet entropy features are extracted from the high frequency sub-bands. The true pedestrian regions are finally classified and recognized using the support vector machine (SVM) classifier. Comparisons between our approach and traditional approaches are presented and experimental results using several thermal infrared image databases show the proposed scheme to be very promising.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Recently, automatic pedestrian detection in thermal infrared imagery for intelligent video surveillance and driver assistance-systems for vehicles has attracted more and more attention [1–4]. However, the problem of detecting pedestrians is very difficult because of the variability in appearance and pose for each pedestrian, especially in outdoor environments. Compared with visible light images, thermal infrared images have unique characteristics. Generally, the target object's intensity is mainly determined by its temperature and radiated heat and is independent of the current light conditions, so the detection system can be applied indiscriminately in both day and night. Also, the infrared images almost eliminate the influences of color, texture, and illumination on the target object's appearance variability. Therefore, there is a latent development potential for pedestrian detection in thermal infrared images. However, infrared images are not perfect either. First, because the target object's surface properties (emissivity, reflectivity, and transmissivity) and wavelength both affect the infrared images intensities, non-human objects and backgrounds, such as animals, cars, light poles, and buildings, produce additional bright areas in infrared images, especially in summer afternoon. These clutters in infrared images make it impossible to accurately detect pedestrians based only on their brightness. Second, due to limitations in camera technology, most infrared images have lower spatial resolution and less sensitivity than visible images which often lead to low image quality, such as blurring, low target-to-background con-

trast, and great noise. So it is a complex challenge to make precise detection of pedestrians using thermal infrared imagery.

In general, pedestrian detection in thermal infrared images is performed in two steps: candidate ROI extraction and pedestrian detection from these candidate regions. The main methods in the first step include template match [5], movement difference [2,6] and intensity oriented projection [7–9]. The template match method requires multiscale searching for the whole image, which is time-consuming. The movement difference method is mainly applied to moving images and is not appropriate for static or single images. This paper therefore uses a statistical adaptive intensity oriented projection method to extract regions of interest. The second step adopts certain pattern recognition or machine learning methods to classify the candidate ROIs and recognize the real pedestrian regions. The key problem is how to describe the pedestrian and discriminate the pedestrians from other candidates based on effective features. Many features are currently used to describe pedestrians, such as shape features (compactness and leanness) [10], histograms of oriented gradients (HOG) [11,12] and shape-independent features (histogram, inertia, and contrast) [9]. This paper uses wavelet entropy features based on the double-density dual-tree complex wavelet transform (DD-DT CWT) to accurately describe the pedestrian.

In this paper, we introduce a complete robust pedestrian detection system for application in thermal infrared images. Based on the self-emission of pedestrians related to the Planck's law, which usually results in the high brightness property of the pedestrian pixels in thermal infrared images, we first adopt the intensity oriented projection method to extract the candidate ROIs which avoids the time-consuming multiscale searching of the whole

* Corresponding author. Tel./fax: +86 23 65112779.

E-mail address: wggong@cqu.edu.cn (W. Gong).

image. We then acquire the wavelet entropy features of candidate ROIs by using DD-DT CWT, which can precisely describe the pedestrian's features. Finally, we use the SVM classifier to recognize the true pedestrians from the candidate regions. Detection results are found for the OTCBVS Benchmark Dataset Collection [13] and comparisons between our approach and traditional detection methods are presented. Experimental results show that the proposed scheme can improve the robustness of the complete system and precisely detect the pedestrians.

2. Candidate pedestrian regions selection

During the selection of candidate pedestrian regions, we use a statistical adaptive intensity oriented projection method which contains horizontal and vertical projections. In order to eliminate the interference from noise and low brightness pixels, we must first segment the original image into a binary image using a flexible threshold T . Considering that the dynamic range of image intensity between frames may vary dramatically due to different seasons or varying outdoor environments, we set the threshold T as a balance between the image maximum intensity I_{max} , the image mean intensity I_{mean} and the pedestrian mean intensity P_{mean} in the training sample set, expressed as

$$T = w_1 I_{max} + w_2 I_{mean} + (1 - w_1 - w_2) P_{mean} \quad (1)$$

where w_1 and w_2 are weights that satisfy $0 \leq w_1, w_2 \leq 1$. The values of w_1 and w_2 are set to acquire the optimal segmentation results through experimental tests. The threshold T will vary adaptively with the dynamic range of image intensity. Then the segmented binary image is projected vertically using the following steps:

- (1) Obtain the intensity vertical projection curves by computing a count of the pixels whose intensities are higher than the threshold T . The intensity vertical projection curves are defined as the number of bright pixels in image columns according to their horizontal positions, as shown in Fig. 1b. Fig. 1a is the original image. The curves can be divided into several heaves or waves with rising left curves and declining right curves.
- (2) Search for each rising left edge and declining right edge from the projection curves.

- (3) Segment the image into several vertical stripes by pairing each rising left edge and declining right edge, as shown in Fig. 1c. Each stripe may contain one or more pedestrians.

Like the intensity vertical projection curve, the horizontal projection curve can also be defined as the number of bright pixels in image rows according to their vertical positions. A horizontal projection is used to acquire horizontal stripes based on the projection curve. Then we can make the bounding boxes from the intersection of the vertical and horizontal stripes. These bounding boxes are called the candidate ROIs where pedestrians may be located.

However, the candidate ROIs may contain some unwanted artefacts due to interference from the many warm targets in the scene. Therefore, an additional intensity oriented projection method is applied to produce more accurate candidate pedestrian ROIs using the following steps:

- (1) Compute the width/height ratio of the candidate ROI as $R_{wh} = l_w / l_h$, where l_w and l_h represent the width and height of the candidate ROI.
- (2) If $R_{wh} \geq T_v$, then the candidate ROI is projected by the second level of vertical oriented projection, where T_v represents the second level of vertical projection threshold.
- (3) Else if $R_{wh} \leq T_h$, then the candidate ROI is projected by the second level of horizontal oriented projection, where T_h represents the second level of horizontal projection threshold.

Fig. 1 shows the results of the intensity oriented projection method for a thermal infrared image using two levels of projections. We can see from Fig. 1d and e that the second level of projection almost eliminates the interference from the trailer phenomenon and acquires more accurate candidate ROIs.

3. Feature extraction

Feature extraction is the key point of the complete pedestrian detection system. The wavelet transform is often used for robust feature extraction from images or time series data for its advantage of local analysis in the spatial and frequency domains. Recently, the dual-tree complex wavelet transform and double-density wavelet transform were both used to extract the mean and variance of

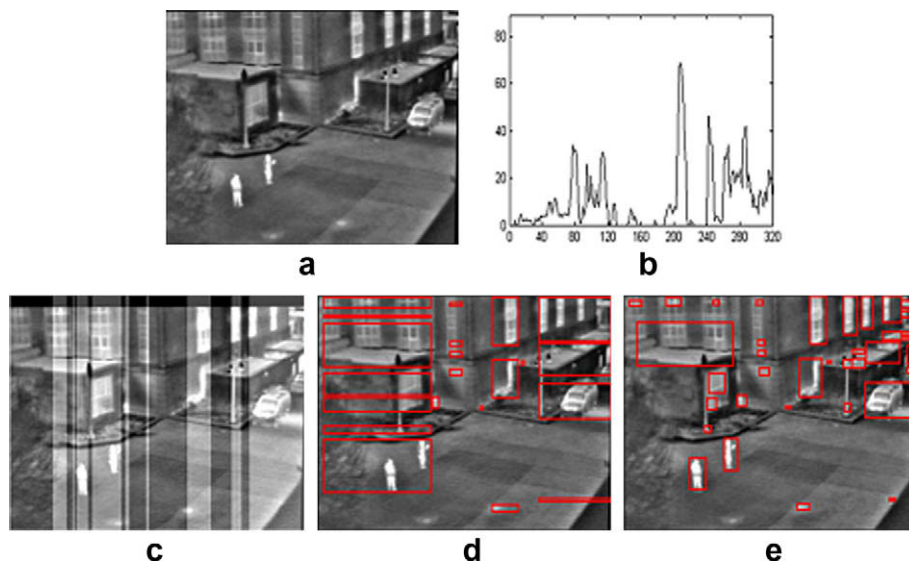


Fig. 1. Results of the intensity oriented projection method for a thermal infrared image: (a) original image, (b) intensity vertical projection curves, (c) horizontal segmentation, (d) first level of projection, and (e) second level of projection.

the wavelet coefficients from each scale and frequency band, and were successfully applied to palmprint recognition [14], texture classification [15], and image retrieval [16,17]. This paper proposes a novel feature extraction method based on DD-DT CWT, which combines both the dual-tree wavelet transform and the double-density wavelet transform. DD-DT CWT describes the pedestrian image with 16 dominant orientations and each orientation is represented using two wavelets, therefore it can improve the accuracy of image decomposition and reconstruction. Also, the wavelet entropy is taken as the feature vector by computing the wavelet coefficients from each transformed frequency band, which reflects the energy distribution in the frequency domain. The final wavelet entropy feature thus reduces the dimensions greatly and therefore improves the speed of the detection system.

3.1. Double-density dual-tree complex wavelet transform

The 2-D dual-tree complex wavelet transform (2-D DT CWT) [18,19] is implemented by using four 2-D discrete wavelet transforms in parallel with different filter banks in rows and columns separately. Twelve wavelets are acquired by taking the sum and difference of each pair of subbands. Six directions are described, each by two wavelets.

The 1-D double-density wavelet transform (1-D DD WT) [20] is based on a scaling function and two different wavelets where one wavelet is a half-sample shift of the other. Compared with traditional wavelet transforms, oversampling is used instead of critical sampling during implementation to reduce the translation sensitivity. The 2-D DD WT [21] uses the 1-D DD WT alternately; the row is filtered first, and then the column. Nine 2-D subbands are created, where one is a 2-D subband and the remaining eight subbands become eight 2-D wavelet filters. The orientation descriptions are like those of the discrete wavelet transform.

The 1-D DD-DT WT [22] combines the 1-D DT WT and the 1-D DD WT. It is constructed using two different scaling functions ($\phi_h(t)$, ϕ_g) and four different wavelets ($\psi_{h,i}$, $\psi_{g,i}$, $i = 1, 2$), where $\psi_{h,1}(t)$ is a half-sample shift of $\psi_{h,2}(t)$ and $\psi_{g,1}(t)$ is a half-sample shift of $\psi_{g,2}(t)$, as follows:

$$\psi_{h,1}(t) = \psi_{h,2}(t - 0.5), \quad \psi_{g,1}(t) = \psi_{g,2}(t - 0.5) \quad (2)$$

The two wavelets become an approximate Hilbert transform pair as

$$\psi_{g,1}(t) = H\{\psi_{h,1}(t)\}, \quad \psi_{g,2}(t) = H\{\psi_{h,2}(t)\} \quad (3)$$

The 2-D DD-DT CWT is implemented using four oversampling 2-D DD DWTs for the input image in parallel. The rows and columns are filtered using different filter banks. Fig. 2 illustrates the

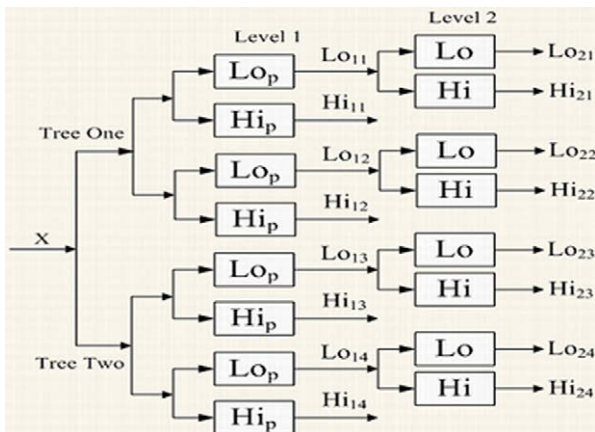


Fig. 2. Two levels of 2D DD-DT CWT.

process of two levels of the 2-D DD-DT CWT. Lo_p and Hi_p are the first level of the decomposition filter banks and represent one scaling filter and eight wavelet filters of 2-D DD DWT. Lo and Hi are the second and remaining levels of decomposition filter banks. Lo_{mn} and Hi_{mn} ($m = 1, 2$; $n = 1, 2, \dots, 4$) represent one low pass subband and eight high pass subbands separately. Therefore, four low pass subbands and 32 high pass subbands are produced for our input image after one level of the transform.

The DD-DT CWT is completed by iteratively transforming the low pass subbands with the determined levels. Finally, 32 wavelets are created by the sum/difference operation on each pair of subbands. Sixteen orientations are described, each by two wavelets.

Fig. 3 shows the comparisons between the 2-D DT CWT, 2-D DD WT and 2-D DD-DT CWT. In Fig. 3a, six directions ($\pm 15^\circ$, $\pm 45^\circ$, $\pm 75^\circ$) are described by 2-D DT CWT using twelve wavelets in the first and second line, each by two wavelets (one for real part, the other for imagery part), and the third line can be taken as the magnitude of the complex wavelet. Fig. 3b illustrates eight wavelets used in 2-D DD WT, the first two wavelets are oriented in the vertical direction, the third and sixth are oriented in the horizontal direction, while the remaining four wavelets have no dominating orientation for the checkerboard affect. Fig. 3c shows 32 wavelets used in 2-D DD-DT CWT, which describe 16 orientations. As we can see, after the 2-D DD-DT CWT, the frequency band in the transform domain is finer, and the orientation information is richer. Also, the 2-D DD-DT CWT eliminates the mosaic phenomenon and acquires more detailed information for the edges of the input image. Therefore we can get more accurate representation for pedestrian regions in thermal infrared images after 2-D DD-DT CWT.

The candidate pedestrian regions are not large in the thermal infrared images, so we set the wavelet decomposition scale factor $j = 3$. Suppose the candidate region size is $m \times n$; we then get three levels of subbands whose sizes are $m/2^j \times n/2^j$. Thirty-two high pass subbands and four low pass subbands are produced after each level of decomposition. Three levels of decompositions are completed by iteratively transforming the low pass subbands. Finally, 96 high pass subbands and four low pass subbands are created in total. Fig. 4 shows the decomposition processes of a pedestrian and car in thermal infrared images using 2-D DD-DT CWT. Two subbands are used to describe the detail information for the same dominant orientation in each level and therefore there are 16 main orientations in all. Each high frequency subband represents the fine-to-coarse description from different orientations and resolutions on the same side of the edges and contours.

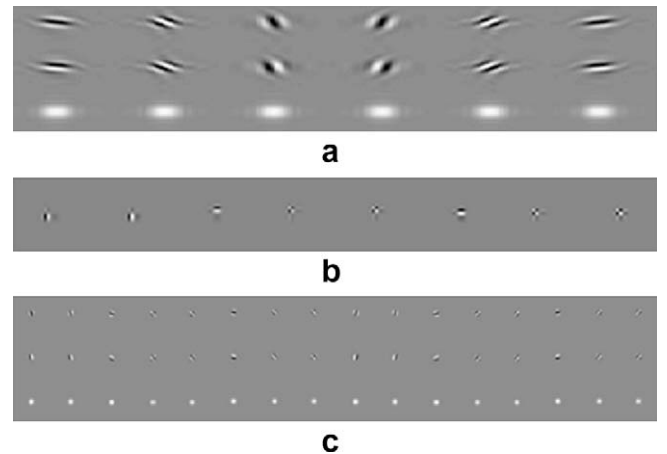


Fig. 3. (a) 2-D DT CWT, (b) 2-D DD WT, and (c) 2-D DD-DT CWT.

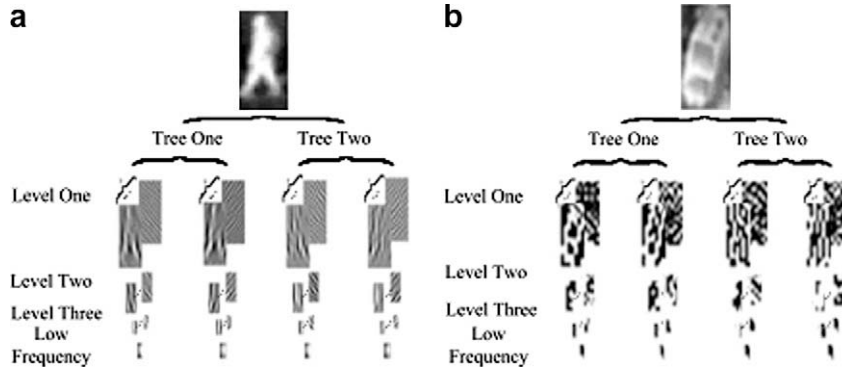


Fig. 4. Decomposition processes of a pedestrian and car using 2-D DD-DT CWT: (a) pedestrian decomposition; (b) car decomposition.

3.2. Wavelet entropy

Wavelet entropy is widely used as a feature extraction method in signal detection and recognition applications. Let the wavelet energy E_j be the quadratic sum of the wavelet coefficients $C_j(k)$ in a single scale j as

$$E_j = \sum_{k=1}^N |C_j(k)|^2 \quad (4)$$

The total energy of the signal is

$$E_{total} = \sum_{j=1}^J E_j = \sum_{j=1}^J \sum_{k=1}^N |C_j(k)|^2 \quad (5)$$

The relative wavelet energy is defined as

$$P_j = \frac{E_j}{E_{total}}, \quad \sum_{j=1}^J P_j = 1 \quad (6)$$

If the wavelet coefficients matrix is taken as a probability distribution series, then the entropy computed by the series reflects the sparseness of the coefficients matrix, which is the order degree of the distribution derived by wavelet entropy [23]. According to the Shannon entropy theory, we define the wavelet entropy as

$$S_w = - \sum_{j=1}^J P_j \ln P_j \quad (7)$$

The wavelet entropy of the pedestrian image reflects the energy distribution of the image in the frequency domain of the wavelet transform. It is defined as

$$H_m(j) = - \sum_{j=1}^J P_{mj} \ln P_{mj} \quad (8)$$

where j is the scale, and m represents the high frequency part. In the 2-D DD-DT CWT, $m = 1, 2 \dots 32$, which includes 16 directions of the real part and 16 directions of the imaginary part. A 1×32 vector is thus produced by extracting the wavelet entropy from the original pedestrian image after 2-D DD-DT CWT, which greatly reduces the feature dimensions.

4. SVM classifier

The support vector machine (SVM) classifier is a binary classification algorithm that tries to find an optimal hyperplane as a decision function in a high dimensional space [24]. It has great generalization ability and is adept at solving pattern recognition problems for non-linear and small sample sets. In this paper, we use a SVM classifier to classify and recognize pedestrians.

In a two-class pattern recognition problem, given a training data set:

$$\{(x_i, y_i); x_i \in R^N; y_i = \pm 1; i = 1, 2 \dots l\} \quad (9)$$

the decision function is set as

$$f(x) = \text{sgn} \left[\sum_{i=1}^l \alpha_i K(x_i, x) + b \right] \quad (10)$$

with the constraint conditions given as:

$$\begin{aligned} \min_{\omega, \xi} & \frac{1}{2} (\omega, \omega) + C \sum_{i=1}^l \xi_i, \quad C \geq 0; \\ y_i [(\omega, x) + b] & \geq 1 - \xi_i, \quad \xi_i \geq 0. \end{aligned} \quad (11)$$

The optimal solution of this problem is then obtained using the Lagrangian theory

$$\max_{\alpha} \sum_{i=1}^l \alpha_i + \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j), \quad 0 \leq \alpha_i \leq C \quad (12)$$

where x_i is the training sample set, y_i is the class label, x is the pending sample, and b is the threshold. $K(\cdot)$ is the kernel function which satisfies the Mercer condition. The role of $K(\cdot)$ is to map the original feature space to high dimensional space and even to infinite dimensional space. The optimal solution of the problem is then sought in the new space. Popular kernel functions include the Gauss kernel, the polynomial kernel, and the Sigmoid kernel.

Training the SVM classifier involves looking for the solutions for a and b that maximize the formula (12). This optimization problem can be tuned by changing the value of the constant C and the parameters of the kernel function. The Gauss kernel function is chosen and the parameter optimized method proposed by Carl

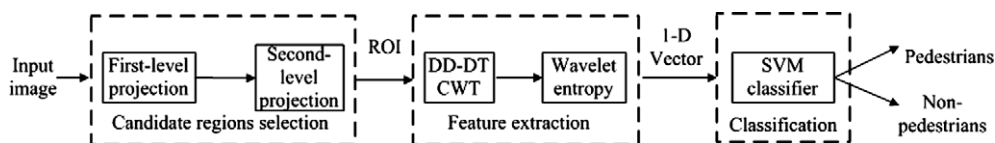
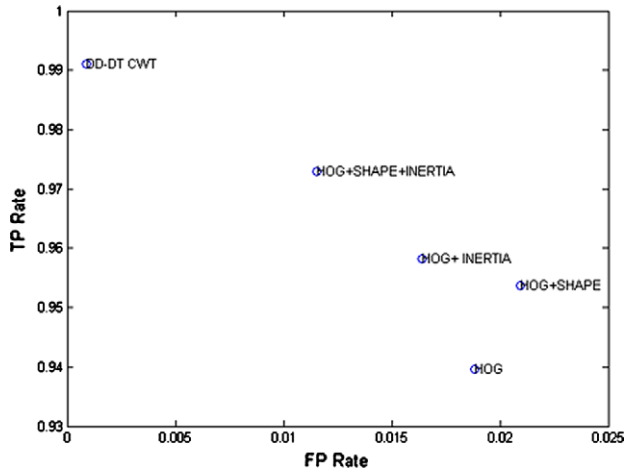
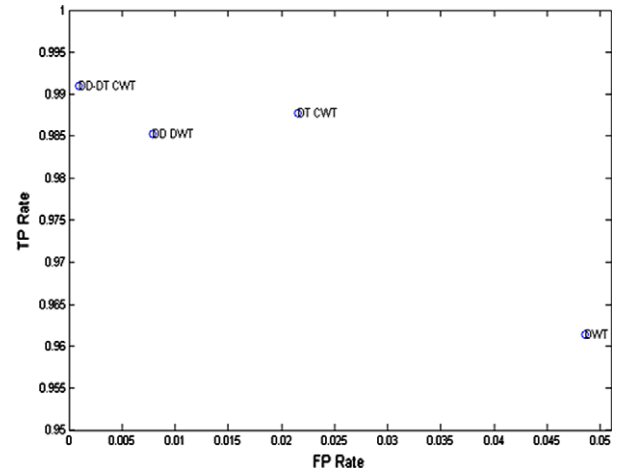
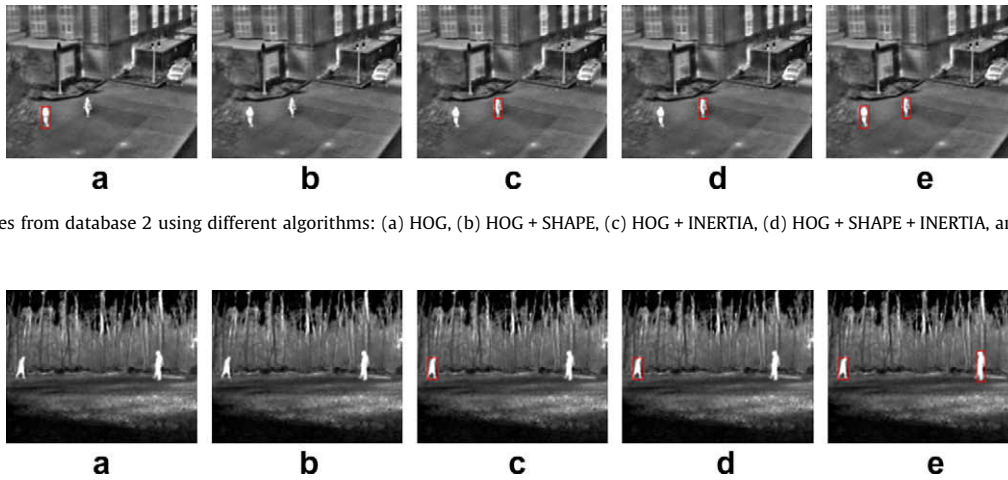


Fig. 5. Block diagram of detection system.

Table 1

Experimental results of different features based pedestrian detection algorithms.

	HOG		HOG + SHAPE		HOG + INERTIA		HOG + SHAPE + INERTIA		WE (DD-DT CWT)	
	FPR (%)	TPR (%)	FPR (%)	TPR (%)	FPR (%)	TPR (%)	FPR (%)	TPR (%)	FPR (%)	TPR (%)
1	12.6	91.0	3.9	94.1	7.9	94.1	3.9	98.1	1.6	98.7
2	1.7	94.4	1.1	95.8	1.8	94.1	1.1	96.4	0	99.2
3	1.3	94.1	2.7	95.1	1.1	97.8	1.0	98.0	0.05	99.0
All	1.9	93.9	2.1	95.4	1.6	95.8	1.2	97.2	0.09	99.1

**Fig. 6.** Comparisons between different algorithms.**Fig. 9.** Comparisons between different wavelet transforms.**Fig. 7.** Test images from database 2 using different algorithms: (a) HOG, (b) HOG + SHAPE, (c) HOG + INERTIA, (d) HOG + SHAPE + INERTIA, and (e) DD-DT CWT.**Fig. 8.** Test images from database 3 using different algorithms: (a) HOG, (b) HOG + SHAPE, (c) HOG + INERTIA, (d) HOG + SHAPE + INERTIA, and (e) DD-DT CWT.

and Peter [25] is used in this paper, and therefore each parameter is tuned adaptively in order to obtain the optimal classification results.

5. Complete system

In this section, we describe a proposal for a complete pedestrian detection system for thermal infrared images. The framework of the system is shown in Fig. 5, which contains three models: the candidate regions selection, feature extraction and classification.

The entire procedure is listed in the following steps:

- (1) Search for all possible candidate ROIs from the input image using the two-level intensity oriented projection method.
- (2) Classify the candidate regions manually, and randomly select 325 pedestrian samples and 336 non-pedestrian samples which are used to construct the training data set.
- (3) Decompose all the training samples by three levels of 2-D DD-DT CWT and therefore obtain 96 high frequency subbands and four low frequency subbands from each training sample.
- (4) Acquire the wavelet entropy by computing the relative energy of each high frequency subband, where a 1×32 vector is created from each training sample.
- (5) Take the 1×32 vector acquired in step 4 as the input feature of the SVM classifier, adaptively tune all of the parameters of the classifier and save the SVM classifier with the best recognition results for the remaining detection process.

Table 2

Experimental results of different wavelet transforms.

	DWT		DT CWT		DD DWT		DD-DT CWT	
	FPR (%)	TPR (%)	FPR (%)	TPR (%)	FPR (%)	TPR (%)	FPR (%)	TPR (%)
1	7.9	92.2	3.1	96.1	3.9	96.1	1.6	98.7
2	6.1	99.0	2.9	99.2	3.8	99.0	0	99.2
3	3.7	94.1	1.6	98.8	0.8	98.6	0.05	99.0
All	4.8	96.1	2.1	98.8	0.8	98.5	0.09	99.1

- (6) Locate the candidate regions of the image awaiting detection, obtain the feature vector using the foregoing methods, and input the computed feature vector into the trained SVM classifier.
- (7) Obtain the classification results and locate the pedestrian regions in the input image.

6. Experimental results

In this section, we report our experimental results with three different thermal infrared image databases from the OTCBVS Benchmark Dataset Collection [13]. The first and second test databases are from the OSU Thermal Pedestrian Database [26] and the OSU Color-Thermal Database [27] separately. The first database contains 284 images with 155 pedestrians in total which were taken at an idle pedestrian intersection on the Ohio State University campus, while the second database contains 400 images with 700 pedestrians which were taken at a busy pedestrian intersection. The third test database is from the Terravic Motion IR Database provided by Roland Mieziako [13], taken

in an outdoor environment, and containing 350 images with 700 pedestrians. All the three databases are divided into train and test sets as usual.

In order to evaluate the most efficient set of parameters w_1 and w_2 listed in Eq. (1), we try different values for each parameter, both from 0.05 to 1 with the step 0.05, to get the flexible threshold T . We then use the different threshold T to segment the train infrared images. It is found that we can get the optimal segmentation results when setting w_1 and w_2 to 0.25 and 0.50 separately. Meanwhile, we acquire the vertical and horizontal projection threshold T_v and T_h according to the statistics of the width/height ratios of 376 pedestrians in the train sample set. In the experiments, the values of T_v and T_h are set to 0.66 and 0.30 separately.

The experiments are divided into two groups. The first group compares the proposed wavelet entropy feature-based detection method with other well-known feature-based detection methods; and the second group compares the performance of different wavelet transforms. We use the True Pedestrians Rate (TPR) and the False Pedestrians Rate (FPR) to estimate the performance of the detection algorithm, where:



Fig. 10. Test images from database 2 using different wavelet transforms: (a) DWT, (b) DT CWT, (c) DD DWT, and (d) DD-DT CWT.

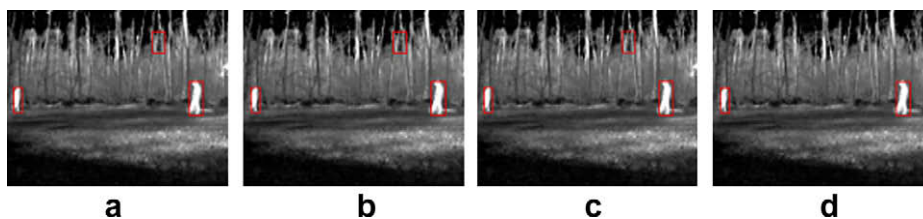


Fig. 11. Test images from database 3 using different wavelet transforms: (a) DWT, (b) DT CWT, (c) DD DWT, and (d) DD-DT CWT.

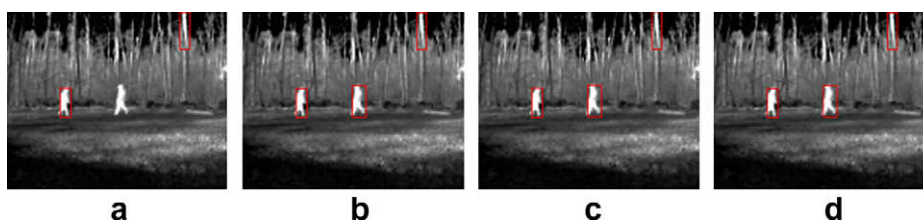


Fig. 12. Test images from database 3 using different wavelet transforms (false alarm case): (a) DWT, (b) DT CWT, (c) DD DWT, and (d) DD-DT CWT.

$$\text{TPR} = \frac{\text{true pedestrians count}}{\text{pedestrians in total}} \quad \text{and}$$

$$\text{FPR} = \frac{\text{false pedestrians count}}{\text{non - pedestrians in total}}$$

6.1. Comparisons between well-known features

The well-known features used at present for pedestrian detection in thermal infrared images are shape features (SHAPE) [10], histograms of oriented gradients (HOG) [11,12], and intensity distribution-based inertia features (INERTIA) [9]. Three thermal infrared image databases are tested with all these features and their potential hybrid features. Table 1 summarizes the experimental results. Fig. 6 illustrates the performance comparisons of the different features based classifiers. From Table 1 and Fig. 6 we can see that the hybrid features using HOG + SHAPE + INERTIA is superior compared to other traditional features, while our proposed wavelet entropy (WE) feature-based classifier is superior compared to hybrid features. The TPR of the WE feature is 1.02 times of the hybrid features, while the FPR of the WE feature is only one thirteenth of the hybrid features. Therefore our proposed feature has the best detection performance. Figs. 7 and 8 show some detection results for real thermal infrared images from test databases 2 and 3 separately (database 1 is ignored because of its simple background and the negligible differences between the experimental results).

6.2. Comparisons between different wavelet transforms

This section presents the comparisons between our proposed DD-DT CWT and other wavelet transforms, such as DWT, DT CWT and DD DWT. The wavelet entropy feature is used in all four wavelet transforms. Table 2 summarizes the experimental results. Fig. 8 shows the comparison of the performances of different wavelet transform based classifiers. We can also see that the proposed DD-DT CWT based classifier has the best detection performance from Table 2 and Figs. 9–12 show some detection results for real thermal infrared images from test databases 2 and 3 separately using the different wavelet transforms.

7. Conclusions

In this paper, a robust method for pedestrian detection in thermal infrared images has been proposed. The main characteristic of this method is the wavelet entropy feature-based classification algorithm using DD-DT CWT. The DD-DT CWT combines the DT CWT and DD DWT and has the advantages of shift invariance, directional selectivity, freedom from aliasing, and a near-continuous wavelet transform. The proposed wavelet entropy properly reflects the energy distribution of the image in the frequency domain of the wavelet transform. The combination of DD-DT CWT and the wavelet entropy can thus describe the image features accurately. Experimental results have shown that our approach is very encouraging. Future work will consider the temporal correlation and motion cues for dealing with image sequences. Moreover, we will combine detection and tracking to improve the performance of our system.

Acknowledgments

This research was supported by the National High-Tech Research and Development Plan of China (863 Program) under Grant

No. 2007AA01Z423, the Foundation Research Project of the 'Eleventh Five-Year-Plan' of China under Grant No. C10020060355, and the Natural Science Foundation Project of CQ CSTC under Grant No. 2008BB2199.

References

- [1] M. Shah, O. Javed, K. Shafique, Automated visual surveillance in realistic scenarios, *IEEE Multimedia* 14 (2007) 30–39.
- [2] A.E. Maadi, X. Maldague, Outdoor infrared video surveillance: a novel dynamic technique for the subtraction of a changing background of IR images, *Infrared Phys. Technol.* 49 (2007) 261–265.
- [3] M. Bertozzi, A. Broggi, A. Fascioli, T. Graf, M. Meinecke, Pedestrian detection for driver assistance using multiresolution infrared vision, *IEEE Trans. Vehicle Technol.* 53 (2004) 1666–1678.
- [4] T. Tsuji, H. Hattori, M. Watanabe, N. Nagaoka, Development of night-vision system, *IEEE Trans. Intell. Transport Syst.* 3 (2002) 203–209.
- [5] H. Nanda, L. Davis, Probabilistic template based pedestrian detection in infrared videos, in: *Proceedings of IEEE Intelligent Vehicle Symposium*, 2002, pp. 15–20.
- [6] J. Davis, V. Sharma, Background-subtraction in thermal imagery using contour saliency, *Int. J. Comput. Vision* 71 (2007) 161–181.
- [7] F.L. Xu, X. Liu, K. Fujimura, Pedestrian detection and tracking with night vision, *IEEE Trans. Intell. Transport Syst.* 6 (2005) 63–71.
- [8] M. Bertozzi, A. Broggi, C. Caraffi, M.D. Rose, M. Felisa, G. Vezzoni, Pedestrian detection by means of far-infrared stereo vision, *Comput. Vision Image Understan.* 106 (2007) 194–204.
- [9] Y.J. Fang, K. Yamada, Y. Ninomiya, B. Horn, I. Masaki, A shape-independent method for pedestrian detection with far-infrared images, *IEEE Trans. Vehicle Technol.* 53 (2004) 1679–1697.
- [10] Congxia Dai, Yunfei Zhengand, Xin Li, Pedestrian detection and tracking in infrared imagery using shape and appearance, *Comput. Vision Image Understan.* 106 (2007) 194–204.
- [11] F. Suard, A. Rakotomamonjy, A. Bensrhair, A. Broggi, Pedestrian detection using infrared images and histograms of oriented gradients, in: *IEEE Intelligent Vehicles Symposium*, 2006, pp. 206–212.
- [12] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, *Int. Conf. Comput. Vision Pattern Recogn.* 2 (2005) 886–893.
- [13] OTCBVS Benchmark Dataset <<http://www.cse.ohio-state.edu/otcbvs-bench>> (accessed 02.09).
- [14] Y.X. Wang, Q.Q. Ruan, X. Pan, Palmprint recognition method using dual-tree complex wavelet transform and local binary pattern histogram, in: *International Symposium on Intelligent Signal Processing and Communication Systems*, 2007, pp. 646–649.
- [15] S. Hatipoglu, S.K. Mitra, N. Kingsbury, Texture classification using dual-tree complex wavelet transform, in: *International Conference on Image Processing and its Applications*, 1999, pp. 344–347.
- [16] A. Mumtaz, S.A.M. Gilani, T. Jameel, A novel color image retrieval system based on dual tree complex wavelet transform and support vector machines, in: *IEEE International Multitopic Conference*, 2006, pp. 163–168.
- [17] R. Peter, N. Kingsbury, Complex wavelets features for fast texture image retrieval, in: *IEEE International Conference on Image Processing*, 1999.
- [18] N.G. Kingsbury, The dual tree complex wavelet transform: a new efficient tool for image restoration and enhancement, in: *European Signal Processing Conference*, 1998, pp. 319–322.
- [19] N.G. Kingsbury, Complex wavelets for shift invariant analysis and filtering of signals, *Appl. Comput. Harmon. Anal.* 10 (2002) 234–253.
- [20] A. Jayawardena, Design of double density wavelet filter banks, *Int. Symp. Signal Process. Appl.* 2 (2003) 463–466.
- [21] J.R. Sveinsson, J.A. Benediktsson, Double density wavelet transformation for speckle reduction of SAR images, *Geosci. Rem. Sens. Symp.* 1 (2002) 113–115.
- [22] I.W. Selesnick, The double-density dual-tree DWT, *IEEE Trans. Signal Process.* 52 (2004) 304–314.
- [23] W.X. Ren, Z.S. Sun, Structural damage identification by using wavelet entropy, *Eng. Struct.* 30 (2008) 2840–2849.
- [24] N. Cristianini, J. Shawe-Taylor, *Introduction to Support Vector Machines*, Cambridge University Press, 2000.
- [25] G. Carl, S. Peter, Model selection for support vector machine classification, *Neurocomputing* 55 (2003) 221–249.
- [26] J. Davis, M. Keck, A two-stage approach to person detection in thermal imagery, *IEEE Workshop Appl. Comput. Vision* 1 (2005) 364–369.
- [27] J. Davis, V. Sharma, Background-subtraction using contour-based fusion of thermal and visible imagery, *Comput. Vision Image Understan.* 106 (2007) 162–182.